

# PREDICTION DE MOBILITE BASEE SUR LA CLASSIFICATION SELON LE PROFIL

CHAMEK LINDA, DAoui MEHAMMED, LALAM MUSTAPHA

## RÉSUMÉ

Les réseaux mobiles de 3<sup>ème</sup> génération permettent aux utilisateurs une mobilité plus étendue et plus souple. Ceux-ci peuvent se déplacer tout en exécutant des applications multimédias et temps réel sur leurs mobiles. Toutefois, plusieurs problèmes sont à résoudre. Parmi ces problèmes nous citons : la localisation, les déconnexions fréquentes et la gestion des ressources.

La prédiction des déplacements peut jouer un grand rôle dans la gestion de la mobilité. Par exemple, la connaissance de la position d'un mobile par le système à l'avance, lui permettra un gain de temps lors de sa recherche (le paging). En effet, le nombre de cellule à pager deviendra limité donc il sera plus facile au réseau de le trouver, et de lui acheminer les appels et les données. La prédiction permet également au réseau de réserver des ressources à l'avance à un mobile dans les futures cellules qu'il va traverser, réduisant ainsi la fréquence des déconnexions.

Nous présentons dans ce papier une solution de prédiction des déplacements des utilisateurs basée sur le datamining, plus précisément la classification des utilisateurs selon leurs profils.

**MOTS CLÉS:** réseau mobile, prédiction de mobilité, datamining, profil, QoS

## 1 INTRODUCTION

Les progrès et les avancées technologiques réalisés ces dernières années dans le domaine de la télécommunication mobile est une réalité qui n'est plus à démontrer. En effet, les nouvelles technologies des réseaux mobiles ouvrent aux utilisateurs des perspectives plus importantes. En plus des communications traditionnelles, ils permettent des transmissions haut débit et temps réel des données. Ils intègrent également d applications multimédia telles que : les jeux interactifs ou la téléphonie IP. Pour pouvoir Ces nouvelles applications exigeantes en qualité de service (QoS), doivent avoir la garantie d'une bonne exécution sans retard, discontinuité ou interruption brusque qui se produisent lors du déplacement du mobile, précisément au moment où il passe d'une cellule à une autre. C'est ce qu'on appelle handoff.

La prédiction de mobilité peut améliorer la QoS en intervenant dans plusieurs fonctions de la gestion de la mobilité telle que la localisation des mobiles dans le réseau. Pour acheminer des appels et des données au mobile, le système envoie des messages de paging dans le réseau pour pouvoir le localiser. Ces messages consomment de la bande passante. Si le réseau peut connaître à l'avance l'itinéraire que va suivre un mobile en cours de son déplacement, il pourra le chercher dans un nombre réduit de cellules, réduisant ainsi le nombre de messages de paging et le temps de recherche.

Dans ce papier, nous présentons une stratégie de prédiction des déplacements basée sur la classification des utilisateurs selon leurs profils. Une fois que le modèle a classé un utilisateur dans une catégorie en fonction de son profil, il examine les anciens déplacements des utilisateurs connus ayant le même profil que lui pour prédire son déplacement futur.

## 2 ETAT DE L'ART

Il existe dans la littérature plusieurs techniques permettant la prédiction de la localisation a priori d'un mobile dans le réseau

L'une des techniques les plus utilisées repose sur la localisation par GPS. Son principe de fonctionnement, est que le mobile envoie la position obtenue par GPS à sa station de base. Cette dernière détermine si le mobile est au bord de sa cellule. A chaque réception de la position du mobile, le système calcule la distance séparant le mobile des cellules voisines, la plus courte (proche) distance est sélectionnée [1]

Dans [2] la solution proposée est basée sur la définition d'un seuil dit de réservation. L'idée est de comparer le signal reçu par le mobile provenant des cellules voisines. Si ce signal est inférieur au seuil, on déduit que le mobile se dirige vers cette cellule. Dans [3], une carte de puissance du signal est maintenue par le système. Elle

représente les différents signaux enregistrés dans différents points de la cellule. On fait appel à cette carte pour savoir la position du mobile, et en extrapoler la position future.

Dans [4], [5], des règles de mobilités sont générées en se basant sur un historique des mouvements que chaque mobile maintient (construit) au cours de son déplacement. Ces règles sont ensuite utilisées dans le processus de prédiction. En effet, il a été observé que les usagers ont tendance à avoir un comportement routinier. Sachant cela, et connaissant le comportement habituel des usagers, il devient possible de déduire la prochaine cellule que va visiter un utilisateur en cours de son déplacement.

Dans [6], les auteurs proposent une technique basée sur l'utilisation d'un réseau de neurone multicouche pour exploiter l'historique des mouvements d'un mobile. Les mouvements récents du mobile sont collectés en premier lieu afin de savoir dans quelle LA (location area) il se trouve. Un modèle de mobilité des usagers est d'abord élaboré, ensuite il sera injecté, comme entrées du réseau de neurone, dans le but de les exploiter et d'en tirer (prédire) la position future du mobile avec la plus grande précision possible. Le modèle de mobilité représente l'historique des mouvements du mobile qui ont été enregistrés dans une durée de temps. Le mouvement est défini en termes de la direction empruntée, et la distance parcourue, et ce dans un intervalle de temps. Le rôle du réseau de neurone est de capturer la relation inconnue entre les valeurs passées et futures du modèle de mobilité ; cela est nécessaire pour la prédiction.

Les auteurs dans [7] proposent un algorithme qui se déroule en trois phases. La première consiste à extraire les mouvements du mobile pour découvrir les régularités des mouvements inter cellulaires, c'est le modèle de mobilité du mobile. Des règles de mobilités sont extraites du modèle précédent dans la 2ème phase. Et enfin, dans la 3ème phase, la prédiction de la mobilité est accomplie en utilisant ces règles.

Dans [8], un nouvel algorithme de prédiction de mobilité est décrit. Dans ce travail, le comportement mobile des utilisateurs est représenté comme la répétition de quelques modèles de mouvements élémentaires. Pour estimer la future position d'un mobile, les auteurs proposent une gestion de mobilité dite agressive PMM (Predictive Mobility Management). Un ensemble d'algorithmes de prévisions de mobilité MMP (Mobile Motion Prediction) est utilisé pour prédire la prochaine position des mobiles et ce en se basant sur leurs historiques de mouvements.

Dans [9], les auteurs proposent un schéma de prédiction combinant entre deux niveaux de prédiction : globale et locale. Le modèle de mobilité globale GMM (Global Mobility Management) est déterminé en termes de cellules traversées par un mobile durant son temps de connexion. Quand au modèle local LMM (Local Mobility Management), il est déterminé en termes d'échantillon de 3-tuples prenant en compte ces trois paramètres : vitesse, direction et position. LMM est utilisé pour modéliser les mouvements intra cellulaires d'un mobile, alors que le GMM est utilisé pour les mouvements inter cellulaires d'un

mobile et ce en associant sa trajectoire actuelle à l'une des règles de mobilité déjà existantes. Cependant, les auteurs ne présentent aucune méthode permettant de découvrir ces règles de mobilité.

Une méthode dite DCP (Dynamic Clustering based Prediction) est présentée dans [10]. Elle est utilisée pour découvrir le modèle de mobilité des usagers à partir d'une collection contenant les trajectoires des mobiles. Ces règles sont ensuite utilisées pour la prédiction. Les trajectoires des usagers rassemblées sont groupées selon leurs similitudes.

Un algorithme adaptatif pour la gestion de mobilité est présenté dans [11]. Il se base sur la construction et la maintenance d'un dictionnaire contenant les mises à jour des trajectoires individuelles des mobiles. L'algorithme proposé peut apprendre les modèles de mobilité des usagers. Néanmoins, il est sensible aux mouvements aléatoires des usagers, et ne peut pas gérer un grand nombre de mobiles à la fois.

Dans [12] et [13] l'auteur présente une technique de prédiction reposant sur la modélisation des déplacements d'un mobile par un système fournis. Ce modèle permet la prédiction en se basant sur les anciens déplacements du mobile et ceux des autres utilisateurs qui vont dans le même sens que lui.

### 3 LE DATAMINING ET LA PRÉDICTION DE MOBILITÉ

Les déplacements des mobiles (usagers) sont souvent engendrés par des besoins socio-économiques et sont régis par la topographie des routes et infrastructures couverts par les différentes cellules composant la zone de localisation tel que : écoles, usines, supermarché, autoroute etc. Les déplacements liés aux besoins socio-économiques sont assez habituels, et par conséquent, présentent un aspect régulier.

Les informations concernant un utilisateur et le caractérisant, autrement dit son profil, sont aussi d'une grande importance. En effet, connaître certains caractères d'un utilisateur, nous aide à savoir où il se trouve avec une grande probabilité. Par exemple, une personne d'un âge compris entre 18ans et 25ans qui est étudiant se trouvera probablement au campus un jour de semaine. Les personnes ayant des revenus font leurs achats dans des boutiques de luxe contrairement aux autres qui préfèrent peut être les supermarchés. Les profils de mobilité de ces personnes sont donc différents. La connaissance du profil des utilisateurs du réseau peut donc nous aider à prédire leurs déplacements. Grouper les mobiles selon leurs profils permet d'exploiter les informations de mobilité des autres mobiles pour prédire la future position d'un utilisateur. Le datamining permet, par sa technique de classification, de regrouper les utilisateurs mobiles dans des catégories selon leurs profils. Le datamining est souvent employé pour désigner l'ensemble des outils permettant l'exploration d'une grande quantité de données, et d'en découvrir des modèles implicites. Ces outils

laissent aux utilisateurs l'initiative de choisir les éléments qu'ils veulent observer ou analyser. [14]

Il existe de nombreuses définitions du datamining tant ce domaine fait l'objet de recherche. Ingénieurs, statisticiens, économistes, concepteurs de logiciels, peuvent avoir des conceptions différentes de ce que recouvre ce terme. Nous retenons une définition qui semble faire le compromis entre différentes conceptions. Nous entendons, donc, par datamining le processus permettant l'extraction d'informations prédictives cachées à partir de large base de données.

Le datamining se déroule en cinq étapes principales présenté dans la figure 1 :

Poser le problème : c'est-à-dire définir les objectifs à atteindre et identifier les individus à qui sont destinés les résultats.

Rechercher et sélectionner les données disponibles et en choisir les variables pertinentes

Préparation des données qui consiste à les nettoyer et les transformer, c'est-à-dire supprimer les données superflues, marginales ; sélectionner es attributs et variables permettant de réduire la taille du problème.

Elaboration du modèle (modélisation) qui constitue le cœur du processus du datamining. Un modèle est la représentation des connaissances sous de formes diverses, et selon le problème à résoudre et la méthode appliquée.

Application du modèle qui est, en fait, l'implantation du modèle élaboré,



Figure 1: Processus du datamining

Il existe plusieurs techniques de datamining. Elles sont complémentaires et sont dédiées à différents problèmes. Les travaux présentés dans [15];[16] résument l'utilisation de ces dernières pour différentes classes de problèmes. Il est à noter qu'il n'y a pas de technique meilleure qu'une autre. Le choix d'une technique plutôt qu'une autre se fait selon le besoin et l'objectif à atteindre. Parmi les tâches traitées par le datamining, nous trouvons :

- La classification et prédiction
- L'estimation
- Le clustering
- Recherche d'association et recherche de séquence
- Séries chronologiques

Et pour la réalisation de ces tâches, plusieurs outils sont utilisées parmi elles

- Les réseaux de neurones
- Les réseaux bayésien
- Les arbres de décision
- Les algorithmes génétiques
- L'algorithme des K-means
- L'algorithme des KNN, ...

## 4 PRÉSENTATION DE LA SOLUTION

### 4.1 Principe

La solution de prédiction que nous proposons s'appuie sur les profils des utilisateurs afin d'extrapoler la cellule future d'un mobile en fonction des anciens déplacements des utilisateurs ayant le même profil. (Figure 2).

On entend par profil, toutes les informations concernant un abonné, tel que : son âge, son lieu de travail, son lieu de résidence, etc. en fait, c'est toutes les informations utiles caractérisant un abonné. L'historique des déplacements des utilisateurs est un facteur tout aussi important à prendre en compte.

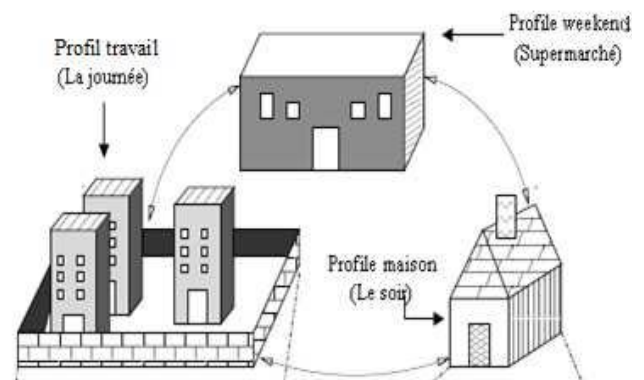


Figure 2: Profil des utilisateurs

Pour mieux illustrer l'importance du profil des utilisateurs et l'historique des mouvements dans le processus de prédiction, prenons, à titre d'exemple, la configuration suivante :

Soit un ensemble de cellules couvrant chacune différentes infrastructures : usine, quartier résidentiel, supermarché, entreprise, boutiques et une cité.

Les utilisateurs en interaction avec ces lieux sont des cadres supérieurs, des employés (simple salarié), leurs familles respectives

Les personnes caractérisées par revenus importants, habitants dans des quartiers résidentiels et possédant des véhicules préfèrent faire leurs achats dans des boutiques de luxes en s'y rendant en utilisant leurs véhicules. Contrairement aux personnes habitants dans une cité et ayant des revenus limités, qui préfèrent prendre le transport en commun pour se rendre au travail ou au supermarché à bas prix pour faire leurs achats.

En se basant sur le profil des utilisateurs, on peut donc extrapoler leurs habitudes et ainsi leurs déplacements.

Si un nouvel utilisateur entre dans le réseau et que nous ne disposons pas de son historique de mouvement (nous ne connaissons pas ses habitudes de déplacements), nous pouvons toujours utiliser l'historique des autres utilisateurs déjà présents dans le réseau, à condition qu'ils aient le même profil que lui. Donc pour pouvoir prédire les déplacements futurs d'un utilisateur (nouveau), il suffit de chercher ses voisins (ayant un profil proche du sien). Sachant que cet utilisateur présente les mêmes caractéristiques que les autres utilisateurs du réseau, nous supposons qu'il adoptera le même profil de déplacement qu'eux.

#### 4.2 Profil des utilisateurs

On entend par profil d'un utilisateur, toutes les informations utiles le caractérisant et permettant au système de comprendre son comportement tel que : son âge, son sexe, son métier, son lieu de travail, son lieu de résidence, ses revenus, etc. ces informations peuvent être récupérées lors de la procédure d'abonnement et enregistrées dans une base de donnée associée au mobile.

#### 4.3 Historisation des Déplacements

Capter l'historique de mobilité d'un utilisateur revient à sauvegarder les différentes transitions effectuées entre les différentes cellules du réseau au cours de son déplacement. Ces informations peuvent être récupérées dans les fichiers log des stations de base gérant la mobilité des utilisateurs.

Chaque cellule maintient un historique des différents déplacements des mobiles ayant visité la Cellule La structure d'une ligne d'historique, inspirée de la structure proposée dans [12], est présentée dans la figure 3 :

Id mobile	Cell source	Cell dest	Moment
-----------	-------------	-----------	--------

Figure 3: Structure d'une ligne d'historique

Avec :

- Id mobile c'est l'identifiant du mobile (unique pour chaque mobile)

- Cell source : indique la cellule d'où vient le mobile -
- Cell destination : indique la cellule où le mobile est allé
- Cell destination : indique la cellule où le mobile est allé
- Moment : indique le type de la journée (férié ou ouvrable)

Ce dernier paramètre est à prendre en considération lors de la capture de l'historique. En effet, les déplacements des mobiles peuvent être très différents en fonction de ce paramètre. Un utilisateur préfère aller dans un lieu de détente le weekend au lieu d'aller au travail.

#### 4.4 Prédiction de localisation

Pour la prédiction de déplacement des mobiles, nous proposons une solution qui se déroule en deux étapes principales : la première étape consiste à classier un nouveau mobile selon son profil, et la deuxième étape (qui est l'étape de prédiction) consiste à exploiter l'historique de ses voisins ayant le même profil que lui pour prédire sa cellule future.

Soit une population de N individus localisés dans une zone géographique couverte par un réseau cellulaire. Chaque individu est caractérisé par un profil défini par : l'âge, le métier, les revenus, lieu de travail, lieu de résidence, le sexe, et la situation familiale...etc. Ces individus utilisent leurs mobiles tout en se déplaçant entre les différentes cellules.

Quand le processus de prédiction est enclenché pour un mobile donné x, notre algorithme de classification calcule les distances le séparant de chaque utilisateur y du réseau. Ce calcul se fera en comparant tous les attributs de leurs profils respectifs en utilisant la distance suivante :

$$D(x; y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Avec :  $x_i$  et  $y_i$  sont les valeurs des attributs de x et y.

Pour garder traces de ces distances, nous utilisons un vecteur tab à N éléments. Nous procédons ensuite au tri de celui-ci ; et la sélection des K plus petites distances correspondantes au K plus proches voisins du mobile.

Une fois cette étape achevée, nous faisons appel à l'historique des déplacements des ces k individus et nous déterminons la cellule de destination la plus fréquente de cet historique qu'on considère comme étant la cellule future du mobile.

L'algorithme proposé est résumé comme suit :

Soit l'ensemble  $I = \{y_1, y_2, \dots, y_n\}$  les N individus se trouvant dans la cellule C

Entrée : x un nouvel individu pour lequel on veut prédire la cellule future

Paramètre k

Sortie : la cellule future à prédire

Algorithme : Pour ( $i = 1$  à N) Faire

- Calculer la distance entre  $y$  et les  $x_i$  de la cellule  $C_d$  ( $x_i, y$ )
- Enregistrer cette distance dans le vecteur  $tab$
- Calculer la distance entre  $y$  et les  $x_i$  de la cellule  $C_d$  ( $x_i, y$ )
- Enregistrer cette distance dans le vecteur  $tab$

Fait

- Trier les distances calculées (le vecteur  $tab$ )
- Sélectionner les  $k$  plus petites distances,
- Sélectionner l'historique des  $k$  individus les plus proches de  $x$

Déterminer la cellule de destination la plus fréquente et la retourner comme la cellule future à prédire. Le déroulement de cet algorithme va nous permettre de classer un nouvel mobile et nous donnera sa cellule destinataire.

## 5 IMPLÉMENTATION D'UN OUTIL D'ÉVALUATION DE LA SOLUTION

Pour évaluer objectivement un algorithme de prédiction, le choix d'un modèle de mobilité réaliste est essentiel. Ce modèle nous permet de reproduire, de façon réaliste le déplacement d'un ensemble d'utilisateurs à l'intérieur d'un réseau. Les travaux présentés dans la littérature utilisent des modèles probabilistes (modèles de Markov, processus poissonien, ...etc.) qui engendrent soit des déplacements trop aléatoires, soit des déplacements trop déterministes. Nous avons opté pour le modèle d'activité présenté dans [17]. Ce modèle est basé sur des travaux effectués par des organismes de planification et utilise des statistiques tirées de cinq années de sondages sur les déplacements réels des usagers. Il simule le déplacement d'un ensemble d'utilisateurs pendant un certain nombre de jours. Les déplacements engendrés reposent sur l'activité de chacun des utilisateurs (travail, étude, etc...), la localisation des lieux de ces activités (maison, lieux de travail écoles,...) ainsi que les chemins qui mènent à ces lieux.

Nous sommes actuellement en phase d'implémentation en java d'un outil d'évaluation de notre solution reposant sur des traces obtenues par le modèle présenté dans [17].

## 6 CONCLUSION

Nous avons présenté dans ce papier notre solution de prédiction de déplacement des mobiles. Notre algorithme se base sur la technique de classification du datamining. La classification des utilisateurs (mobiles) est faite en tenant compte de leurs profils individuels et personnels. La simulation de notre solution est en cours de réalisation en utilisant un simulateur de mouvement réaliste.

## REFERENCES

- [1] I. Feathertone, N. Zhang "Amobility monitoring based advance reservation protocol" Q2Swin et 06 spin Oct 2006
- [2] CK Chua, SM Jiang "Measurement based dynamic bandwidth reservation scheme for handoff in mobile multimedia networks"; conference on universal personalcommunication Oct 98
- [3] S Das, R Jayaram, N Kakani "A call admission & control scheme for QOS provisioning in next generation wireless networks"; wireless network 6 2000
- [4] Vut Hong Nhan, K H Ryu "Future location prediction of moving objects based on movements rules"; LNCIS 344 ICIC 2006
- [5] T Liu, P Bahl "Mobility modeling location tracking and trajectory prediction in wireless ATM networks"; IEEE vol °16 n°06 Aug 98
- [6] B P Vijay Kumar, P Venkataram "Prediction based location management using multilayer neural networks" ; Indian institute of science 2002
- [7] G Yavas, D.Katsaros, O Ulusoy, Y Manolopoulos "A datamining approach for location prediction in mobile environments" ; ELSEVIER 2004
- [8] G.Y.Liu, G.Q.Maguire "A predictive mobility management algorithm for wireless mobile computing and communication"; royal institute of technology 1995
- [9] T.Liu, P. Bahl, Imrich Chlamtac "Mobility modeling, location tracking and trajectory prediction in wireless ATM network"; IEEE 1998
- [10] D. Katsaros, A. Nanopoulos, M. Karakaya, G. Yavas, O. Ulusoy, Y. Manolopoulos "Clustering mobile trajectories for resource allocation in mobile environments" Springer verlay 2003
- [11] A. Bhattacharya, S. K. Das "Le zi update: an information theoretic approach to track mobile users in PCS networks" ACM wireless network 2002
- [12] M. Daoui "Reservation de ressource et prediction de cellule dans un réseau mobile 3G" thèse doctorat 2009
- [13] M.Daoui, A.M'Zoughi, M.Lalam, R. Aoudjid, M. Belkadi "Forcasting models, methods and applications, mobility prediction in cellular networks". I-concepts press 2010
- [14] G Helou, C Aboukhalil "Datamining: techniques d'extraction des connaissances" Université Panthéon Assas paris II 2004
- [15] Paul Balez, Mathieu Beal "Algorithme du datamining" ; EPITA Fev 2002
- [16] "Datamining and applications in genomics, chapter 2 datamining algorithm" 2008.
- [17] J. Scourias and T. Kunz, An Activity-based Mobility Model and Location Management Simulation Framework, Proc., Second ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM), 1999, PP. 61-68.